

SEQUENCE LISTING

<110> van Rooijen, Gijs
Keon, Richard Glenn
Boothe, Joseph
Shen, Yin

<120> Commercial Production of Chymosin in Plants

<130> 9369-153

<140>
<141>

<160> 4

<170> PatentIn Ver. 2.0

<210> 1
<211> 1173
<212> DNA
<213> Bovine

<220>
<221> CDS
<222> (1)...(1173)

<400> 1
atg aac ttc ctt aag tct ttc cct tac gct ttc ctt tgt ttc ggt 48
Met Asn Phe Leu Lys Ser Phe Pro Phe Tyr Ala Phe Leu Cys Phe Gly
 1 5 10 15

caa tac ttc gtt gct gtt act cac gct gag atc acc cgc att cct 96
Gln Tyr Phe Val Ala Val Thr His Ala Ala Glu Ile Thr Arg Ile Pro
 20 25 30

ctc tac aaa ggt aag tct ctc cgt aag gcg ctg aag gaa cat gga ctt 144
Leu Tyr Lys Gly Lys Ser Leu Arg Lys Ala Leu Lys Glu His Gly Leu
 35 40 45

cta gaa gac ttc ttg cag aaa caa cag tat ggc atc agc agc aag tac 192
Leu Glu Asp Phe Leu Gln Lys Gln Gln Tyr Gly Ile Ser Ser Lys Tyr
 50 55 60

tcc ggc ttc ggt gaa gtt gct agc gtg cca ctt acc aac tac ctt gat 240
Ser Gly Phe Gly Glu Val Ala Ser Val Pro Leu Thr Asn Tyr Leu Asp
 65 70 75 80

agt caa tac ttt ggg aag atc tac ctc gga acc ccg cct caa gag ttc 288
Ser Gln Tyr Phe Gly Lys Ile Tyr Leu Gly Thr Pro Pro Gln Glu Phe
 85 90 95

acc gtt ctc ttt gat act ggt tcc tct gac ttc tgg gtt ccc tct atc 336
Thr Val Leu Phe Asp Thr Gly Ser Ser Asp Phe Trp Val Pro Ser Ile
 100 105 110

tac tgc aag agc aat gcc tgc aag aac cac caa aga ttc gat ccg aga 384
Tyr Cys Lys Ser Asn Ala Cys Lys Asn His Gln Arg Phe Asp Pro Arg
 115 120 125

aag tcg tcc acc ttc cag aac tta ggc aaa ccc ttg tct ata cac tac 432
Lys Ser Ser Thr Phe Gln Asn Leu Gly Lys Pro Leu Ser Ile His Tyr
 130 135 140

ggt aca ggt agc atg caa gga atc tta ggc tat gat acc gtc act gtc 480

Gly Thr Gly Ser Met Gln Gly Ile Leu Gly Tyr Asp Thr Val Thr Val		
145 150 155 160		
tcc aac att gtg gac att caa cag aca gta gga ctt agc acc caa gaa	528	
Ser Asn Ile Val Asp Ile Gln Gln Thr Val Gly Leu Ser Thr Gln Glu		
165 170 175		
cca ggt gat gtc ttc acc tat gca gaa ttc gat ggc atc ctt ggt atg	576	
Pro Gly Asp Val Phe Thr Tyr Ala Glu Phe Asp Gly Ile Leu Gly Met		
180 185 190		
gca tac cca tcg ctc gcg tca gag tac tcg ata cct gtg ttt gac aac	624	
Ala Tyr Pro Ser Leu Ala Ser Glu Tyr Ser Ile Pro Val Phe Asp Asn		
195 200 205		
atg atg aac cga cac cta gta gct caa gac ttg ttc tcg gtt tac atg	672	
Met Met Asn Arg His Leu Val Ala Gln Asp Leu Phe Ser Val Tyr Met		
210 215 220		
gac agg aat ggc cag gag agc atg ctc acg ctt gga gct att gat cca	720	
Asp Arg Asn Gly Gln Glu Ser Met Leu Thr Leu Gly Ala Ile Asp Pro		
225 230 235 240		
tcc tac tac aca gga tct ctt cac tgg gtt cca gtc act gtg cag cag	768	
Ser Tyr Tyr Thr Gly Ser Leu His Trp Val Pro Val Thr Val Gln Gln		
245 250 255		
tac tgg caa ttc act gtg gac agt gtc acc atc agc ggt gtg gtt gtt	816	
Tyr Trp Gln Phe Thr Val Asp Ser Val Thr Ile Ser Gly Val Val Val		
260 265 270		
gca tgt gaa ggt gga tgt caa gct atc ttg gat acc ggt acg tcc aag	864	
Ala Cys Glu Gly Cys Gln Ala Ile Leu Asp Thr Gly Thr Ser Lys		
275 280 285		
ctg gtc gga cct agc agc gac att ctc aac att cag caa gct att gga	912	
Leu Val Gly Pro Ser Ser Asp Ile Leu Asn Ile Gln Gln Ala Ile Gly		
290 295 300		
gcc aca cag aac cag tac ggt gag ttt gac ata gat tgc gac aac ctt	960	
Ala Thr Gln Asn Gln Tyr Gly Glu Phe Asp Ile Asp Cys Asp Asn Leu		
305 310 315 320		
agc tac atg cct aca gtt gtc ttt gag atc aac ggc aag atg tac cca	1008	
Ser Tyr Met Pro Thr Val Val Phe Glu Ile Asn Gly Lys Met Tyr Pro		
325 330 335		
ctg acc ccc tcc gcc tat acc agc cag gat caa ggg ttc tgc acc agt	1056	
Leu Thr Pro Ser Ala Tyr Thr Ser Gln Asp Gln Gly Phe Cys Thr Ser		
340 345 350		
gga ttc cag agt gag aac cat tcc cag aaa tgg atc ttg gga gat gtg	1104	
Gly Phe Gln Ser Glu Asn His Ser Gln Lys Trp Ile Leu Gly Asp Val		
355 360 365		
ttc att cgt gag tac tac agc gtc ttt gac agg gcc aac aac ctc gtt	1152	
Phe Ile Arg Glu Tyr Tyr Ser Val Phe Asp Arg Ala Asn Asn Leu Val		
370 375 380		
ggg cta gct aaa gca atc tga	1173	
Gly Leu Ala Lys Ala Ile		
385 390		

<210> 2
<211> 390
<212> PRT

<213> Bovine

<400> 2

Met Asn Phe Leu Lys Ser Phe Pro Phe Tyr Ala Phe Leu Cys Phe Gly
1 5 10 15

Gln Tyr Phe Val Ala Val Thr His Ala Ala Glu Ile Thr Arg Ile Pro
20 25 30

Leu Tyr Lys Gly Lys Ser Leu Arg Lys Ala Leu Lys Glu His Gly Leu
35 40 45

Leu Glu Asp Phe Leu Gln Lys Gln Gln Tyr Gly Ile Ser Ser Lys Tyr
50 55 60

Ser Gly Phe Gly Glu Val Ala Ser Val Pro Leu Thr Asn Tyr Leu Asp
65 70 75 80

Ser Gln Tyr Phe Gly Lys Ile Tyr Leu Gly Thr Pro Pro Gln Glu Phe
85 90 95

Thr Val Leu Phe Asp Thr Gly Ser Ser Asp Phe Trp Val Pro Ser Ile
100 105 110

Tyr Cys Lys Ser Asn Ala Cys Lys Asn His Gln Arg Phe Asp Pro Arg
115 120 125

Lys Ser Ser Thr Phe Gln Asn Leu Gly Lys Pro Leu Ser Ile His Tyr
130 135 140

Gly Thr Gly Ser Met Gln Gly Ile Leu Gly Tyr Asp Thr Val Thr Val
145 150 155 160

Ser Asn Ile Val Asp Ile Gln Gln Thr Val Gly Leu Ser Thr Gln Glu
165 170 175

Pro Gly Asp Val Phe Thr Tyr Ala Glu Phe Asp Gly Ile Leu Gly Met
180 185 190

Ala Tyr Pro Ser Leu Ala Ser Glu Tyr Ser Ile Pro Val Phe Asp Asn
195 200 205

Met Met Asn Arg His Leu Val Ala Gln Asp Leu Phe Ser Val Tyr Met
210 215 220

Asp Arg Asn Gly Gln Glu Ser Met Leu Thr Leu Gly Ala Ile Asp Pro
225 230 235 240

Ser Tyr Tyr Thr Gly Ser Leu His Trp Val Pro Val Thr Val Gln Gln
245 250 255

Tyr Trp Gln Phe Thr Val Asp Ser Val Thr Ile Ser Gly Val Val Val
260 265 270

Ala Cys Glu Gly Gly Cys Gln Ala Ile Leu Asp Thr Gly Thr Ser Lys
275 280 285

Leu Val Gly Pro Ser Ser Asp Ile Leu Asn Ile Gln Gln Ala Ile Gly
290 295 300

Ala Thr Gln Asn Gln Tyr Gly Glu Phe Asp Ile Asp Cys Asp Asn Leu
305 310 315 320

Ser Tyr Met Pro Thr Val Val Phe Glu Ile Asn Gly Lys Met Tyr Pro
325 330 335

Leu Thr Pro Ser Ala Tyr Thr Ser Gln Asp Gln Gly Phe Cys Thr Ser

340

345

350

Gly Phe Gln Ser Glu Asn His Ser Gln Lys Trp Ile Leu Gly Asp Val
355 360 365

Phe Ile Arg Glu Tyr Tyr Ser Val Phe Asp Arg Ala Asn Asn Leu Val
370 375 380

Gly Leu Ala Lys Ala Ile
385 390

<210> 3
<211> 3957
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1554)..(2726)

<220>
<223> Description of Artificial Sequence: Figure 2

<400> 3
ctgcaggaat tcattgtact cccagtatca ttatagtcaa agttttggct ctctcgccgg 60
tggttttta cctctattta aaggggtttt ccacctaaaa attctggtat cattctcact 120
ttacttgtta cttaatttc tcataatctt tggttgaat tatacgcctt ccgcacacga 180
tatccctaca aatttattat ttgttaaaca ttttcaaacc gcataaaatt ttatgaagtc 240
ccgtctatct ttaatgttagt ctaacattt catattgaaa tatataattt acttaattt 300
agcggttgta gaaagcataa agatttattc ttattcttct tcataataat gtttaatata 360
caatataaac aaattcttta ccttaagaag gatttcccat tttatattt aaaaatatata 420
ttatcaaata ttttcaacc acgtaaatct cataataata agttgttca aaagtaataa 480
aatttaactc cataattttt ttattcgact gatcttaaag caacacccag tgacacaact 540
agccattttt ttctttgaat aaaaaaatcc aattatcatt gtatttttt tatacatga 600
aaatttcacc aaacaatcat ttgtggatt tctgaagcaa gtcatgttat gcaaaattct 660
ataattccca tttgacacta cggaagtaac tgaagatctg cttaatcatg cgagacacat 720
cttctaaagt aatttaata atagttacta tattcaagat ttcatatatac aaataactcaa 780
tattacttct aaaaaattaa ttagatataa ttaaaatatt actttttaa tttaagtt 840
aattgttcaa tttgtgacta ttgatttatt attctactat gtttaaattt gttttagat 900
agtttaaagt aaatataagt aatgttagtag agtgttagag tgtaacccta aaccataaac 960
tataacattt atggggact aatttcata tatttcttac tgctttacc ttttcttgg 1020
atgtaagtcc gtaactagaa ttacagtggg ttgccatggc actctgtggc cttttgggtc 1080
atgcatgggt cttgcgcaag aaaaagacaa agaacaaga aaaaagacaa aacagagaga 1140
caaaacgcaa tcacacaacc aactcaaatt agtcaactggc tgatcaagat cgccgcgtcc 1200
atgtatgtct aatgccatg caaagcaaca cgtgcttaac atgcacttta aatggctcac 1260

ccatctcaac ccacacacaa acacattgcc ttttcttca tcatcaccac aaccacctgt 1320
atatattcat tctcttccgc cacctaatt tcttcacttc aacacacgtc aacctgcata 1380
tgcggtgtcat cccatgccc aatctccatg catgttccaa ccaccttctc tcttatataa 1440
tacctataaa tacctcta atcactcact tctttcatca tccatccatc cagagtacta 1500
ctactctact actataatac cccaaaccaa ctcataattca atactactct act atg 1556
Met 1

aac ttc ctt aag tct ttc cct ttc tac gct ttc ctt tgt ttc ggt caa 1604
Asn Phe Leu Lys Ser Phe Pro Phe Tyr Ala Phe Leu Cys Phe Gly Gln
5 10 15

tac ttc gtt gct gtt act cac gct gct gag atc acc cgc att cct ctc 1652
Tyr Phe Val Ala Val Thr His Ala Ala Glu Ile Thr Arg Ile Pro Leu
20 25 30

tac aaa ggt aag tct ctc cgt aag gcg ctg aag gaa cat gga ctt cta 1700
Tyr Lys Gly Lys Ser Leu Arg Lys Ala Leu Lys Glu His Gly Leu Leu
35 40 45

gaa gac ttc ttg cag aaa caa cag tat ggc atc agc agc aag tac tcc 1748
Glu Asp Phe Leu Gln Lys Gln Gln Tyr Gly Ile Ser Ser Lys Tyr Ser
50 55 60 65

ggc ttc ggt gaa gtt gct agc gtg cca ctt acc aac tac ctt gat agt 1796
Gly Phe Gly Glu Val Ala Ser Val Pro Leu Thr Asn Tyr Leu Asp Ser
70 75 80

caa tac ttt ggg aag atc tac ctc gga acc ccg cct caa gag ttc acc 1844
Gln Tyr Phe Gly Lys Ile Tyr Leu Gly Thr Pro Pro Gln Glu Phe Thr
85 90 95

gtt ctc ttt gat act ggt tcc tct gac ttc tgg gtt ccc tct atc tac 1892
Val Leu Phe Asp Thr Gly Ser Ser Asp Phe Trp Val Pro Ser Ile Tyr
100 105 110

tgc aag agc aat gcc tgc aag aac cac caa aga ttc gat ccg aga aag 1940
Cys Lys Ser Asn Ala Cys Lys Asn His Gln Arg Phe Asp Pro Arg Lys
115 120 125

tcg tcc acc ttc cag aac tta ggc aaa ccc ttg tct ata cac tac ggt 1988
Ser Ser Thr Phe Gln Asn Leu Gly Lys Pro Leu Ser Ile His Tyr Gly
130 135 140 145

aca ggt agc atg caa gga atc tta ggc tat gat acc gtc act gtc tcc 2036
Thr Gly Ser Met Gln Gly Ile Leu Gly Tyr Asp Thr Val Thr Val Ser
150 155 160

aac att gtg gac att caa cag aca gta gga ctt agc acc caa gaa cca 2084
Asn Ile Val Asp Ile Gln Gln Thr Val Gly Leu Ser Thr Gln Glu Pro
165 170 175

ggt gat gtc ttc acc tat gca gaa ttc gat ggc atc ctt ggt atg gca 2132
Gly Asp Val Phe Thr Tyr Ala Glu Phe Asp Gly Ile Leu Gly Met Ala
180 185 190

tac cca tcg ctc gcg tca gag tac tcg ata cct gtg ttt gac aac atg 2180
Tyr Pro Ser Leu Ala Ser Glu Tyr Ser Ile Pro Val Phe Asp Asn Met
195 200 205

atg aac cga cac cta gta gct caa gac ttg ttc tcg gtt tac atg gac 2228
Met Asn Arg His Leu Val Ala Gln Asp Leu Phe Ser Val Tyr Met Asp
210 215 220 225

agg aat ggc cag gag agc atg ctc acg ctt gga gct att gat cca tcc Arg Asn Gly Gln Glu Ser Met Leu Thr Leu Gly Ala Ile Asp Pro Ser 230 235 240	2276
tac tac aca gga tct ctt cac tgg gtt cca gtc act gtg cag cag tac Tyr Tyr Thr Gly Ser Leu His Trp Val Pro Val Thr Val Gln Gln Tyr 245 250 255	2324
tgg caa ttc act gtg gac agt gtc acc atc agc ggt gtg gtt gtt gca Trp Gln Phe Thr Val Asp Ser Val Thr Ile Ser Gly Val Val Val Ala 260 265 270	2372
tgt gaa ggt gga tgt caa gct atc ttg gat acc ggt acg tcc aag ctg Cys Glu Gly Gly Cys Gln Ala Ile Leu Asp Thr Gly Thr Ser Lys Leu 275 280 285	2420
gtc gga cct agc agc gac att ctc aac att cag caa gct att gga gcc Val Gly Pro Ser Ser Asp Ile Leu Asn Ile Gln Gln Ala Ile Gly Ala 290 295 300 305	2468
aca cag aac cag tac ggt gag ttt gac ata gat tgc gac aac ctt agc Thr Gln Asn Gln Tyr Gly Glu Phe Asp Ile Asp Cys Asp Asn Leu Ser 310 315 320	2516
tac atg cct aca gtt gtc ttt gag atc aac ggc aag atg tac cca ctg Tyr Met Pro Thr Val Val Phe Glu Ile Asn Gly Lys Met Tyr Pro Leu 325 330 335	2564
acc ccc tcc gcc tat acc agc cag gat caa ggg ttc tgc acc agt gga Thr Pro Ser Ala Tyr Thr Ser Gln Asp Gln Gly Phe Cys Thr Ser Gly 340 345 350	2612
ttc cag agt gag aac cat tcc cag aaa tgg atc ttg gga gat gtg ttc Phe Gln Ser Glu Asn His Ser Gln Lys Trp Ile Leu Gly Asp Val Phe 355 360 365	2660
att cgt gag tac tac agc gtc ttt gac agg gcc aac aac ctc gtt ggg Ile Arg Glu Tyr Tyr Ser Val Phe Asp Arg Ala Asn Asn Leu Val Gly 370 375 380 385	2708
cta gct aaa gca atc tga agcttaataa gtatgaacta aaatgcattgt Leu Ala Lys Ala Ile 390	2756
aggtgttaaga gctcatggag agcatgaaat attgttatccg accatgttaac agtataataa ctgagctcca tctcacttct tctatgaata aacaaaggat gttatgatat attaacactc tatctatgca ccttattgtt ctatgataaa ttccctctta ttattataaa tcattctaat cgtgacggct tatggaatgc ttcaaatagt acaaaaacaa atgtgtacta taagactttc taaacaattc taacttttagc attgtgaacg agacataagt gtttggaaataa cataacaatt ataatggaag aagtttgct ccatttatatt attatataatt acccacttat gtattatatt aggatgttaa ggagacataa caattataaa gagagaagtt tgtatccatt tatataatt atactaccca ttttatattt atacttatcc acttattaa tgtcttata aggtttgatc catgatattt ctaatattt agttgatatg tatatgaaag ggtactattt gaactctctt actctgtata aaggttggat catccttaaa gtgggtctat ttaattttat tgcttcttac agataaaaaaa aaaattatga gttggttga taaaatattg aaggatttaa aataataata aataataat aacatataat atatgttat atatgttat aatatttata taatataaca tttatctata	2816 2876 2936 2996 3056 3116 3176 3236 3296 3356 3416 3476

aaaaagtaaa tattgtcata aatctataca atcgtttagc cttgctggac gactctcaat 3536
tatttaaacg agagtaaaca tatggactt ttgggttatt taacaatta ttatccaaca 3596
ctatatgaaa tttttttt ttatcgccaa ggaaataaaa ttaaatttagg agggacaatg 3656
gtgtgtccca atccctatac aaccaacttc cacaggaagg tcaggtcggg gacaacaaa 3716
aaacaggcaa gggaaatttt ttaatttggg ttgtcttgtt tgctgcataa tttatgcagt 3776
aaaacactac acataaccct ttttagcagta gagcaatggg tgaccgtgtg cttagcttct 3836
tttattttat ttgttatca gcaaagaata aataaaataa aatgagacac ttcaaggatg 3896
tttcaaccct tatacaaaaac cccaaaaaca agtttcctag caccctacca actaaggtac 3956
C 3957

<210> 4
<211> 390
<212> PRT
<213> Artificial Sequence

<400> 4
Met Asn Phe Leu Lys Ser Phe Pro Phe Tyr Ala Phe Leu Cys Phe Gly
1 5 10 15
Gln Tyr Phe Val Ala Val Thr His Ala Ala Glu Ile Thr Arg Ile Pro
20 25 30
Leu Tyr Lys Gly Lys Ser Leu Arg Lys Ala Leu Lys Glu His Gly Leu
35 40 45
Leu Glu Asp Phe Leu Gln Lys Gln Tyr Gly Ile Ser Ser Lys Tyr
50 55 60
Ser Gly Phe Gly Glu Val Ala Ser Val Pro Leu Thr Asn Tyr Leu Asp
65 70 75 80
Ser Gln Tyr Phe Gly Lys Ile Tyr Leu Gly Thr Pro Pro Gln Glu Phe
85 90 95
Thr Val Leu Phe Asp Thr Gly Ser Ser Asp Phe Trp Val Pro Ser Ile
100 105 110
Tyr Cys Lys Ser Asn Ala Cys Lys Asn His Gln Arg Phe Asp Pro Arg
115 120 125
Lys Ser Ser Thr Phe Gln Asn Leu Gly Lys Pro Leu Ser Ile His Tyr
130 135 140
Gly Thr Gly Ser Met Gln Gly Ile Leu Gly Tyr Asp Thr Val Thr Val
145 150 155 160
Ser Asn Ile Val Asp Ile Gln Gln Thr Val Gly Leu Ser Thr Gln Glu
165 170 175
Pro Gly Asp Val Phe Thr Tyr Ala Glu Phe Asp Gly Ile Leu Gly Met
180 185 190
Ala Tyr Pro Ser Leu Ala Ser Glu Tyr Ser Ile Pro Val Phe Asp Asn
195 200 205
Met Met Asn Arg His Leu Val Ala Gln Asp Leu Phe Ser Val Tyr Met
210 215 220

Asp Arg Asn Gly Gln Glu Ser Met Leu Thr Leu Gly Ala Ile Asp Pro
225 230 235 240

Ser Tyr Tyr Thr Gly Ser Leu His Trp Val Pro Val Thr Val Gln Gln
245 250 255

Tyr Trp Gln Phe Thr Val Asp Ser Val Thr Ile Ser Gly Val Val Val
260 265 270

Ala Cys Glu Gly Gly Cys Gln Ala Ile Leu Asp Thr Gly Thr Ser Lys
275 280 285

Leu Val Gly Pro Ser Ser Asp Ile Leu Asn Ile Gln Gln Ala Ile Gly
290 295 300

Ala Thr Gln Asn Gln Tyr Gly Glu Phe Asp Ile Asp Cys Asp Asn Leu
305 310 315 320

Ser Tyr Met Pro Thr Val Val Phe Glu Ile Asn Gly Lys Met Tyr Pro
325 330 335

Leu Thr Pro Ser Ala Tyr Thr Ser Gln Asp Gln Gly Phe Cys Thr Ser
340 345 350

Gly Phe Gln Ser Glu Asn His Ser Gln Lys Trp Ile Leu Gly Asp Val
355 360 365

Phe Ile Arg Glu Tyr Tyr Ser Val Phe Asp Arg Ala Asn Asn Leu Val
370 375 380

Gly Leu Ala Lys Ala Ile
385 390